



ATTORNEY'S DOCKET NO. 96-108

PATENTS

UNITED STATES PATENT APPLICATION

OF

EITAN BACHMAT

FOR

SYSTEM AND METHOD FOR ASSESSING AND CONTROLLING THE OPERATIONAL
EFFECTIVENESS OF A CACHE MEMORY IN A DIGITAL DATA PROCESSING SYSTEM

Certificate of Express Mailing

Express Mail Mailing Label No. EH 629 006 160 US

Date of Deposit April 4, 1997

I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office To Addressee" Service under 37 CFR 1.10 on the date indicated above and is addressed to the Commissioner of Patents and Trademarks, Washington, D. C. 20231.

By Richard A. Jordan

Richard A. Jordan

RECEIVED



1 **FIELD OF THE INVENTION**

2 The invention relates generally to the field of digital data processing systems, and more
3 particularly to a system and method for assessing the operational effectiveness of a cache memory in
4 a digital data processing system. The cache memory operational effectiveness assessment can be used
5 to determine whether increasing the size of a cache memory would provide any significant increase
6 in processing efficiency by the digital data processing system, as well as whether any significant
7 decrease in processing efficiency might occur if the size of the cache memory is decreased.

8 **BACKGROUND OF THE INVENTION**

9 Digital data processing systems include one or more processors for performing processing
10 operations in connection with information stored in a memory. Typically, a memory in a modern
11 digital data processing system consists of a hierarchy of storage elements, extending from large-
12 capacity but relatively slow storage elements and various levels of lower-capacity and relatively fast
13 storage devices. The large-capacity and relatively slow devices include such types of devices as disk
14 or tape storage devices which store information on a magnetic medium; such devices are relatively
15 inexpensive on a storage cost per unit of storage basis. Intermediate in the hierarchy, both in terms
16 of speed and storage capacity are random-access memories, which are somewhat faster than the disk
17 or tape devices, but which are also more expensive on a storage cost per unit of storage basis. At the
18 fastest end of the hierarchy are cache memories, which are also the most expensive and thus generally
19 the smallest.

20 Generally, during processing operations, a processor will enable information to be processed
21 to be copied from the slower devices to the increasingly faster devices for faster retrieval. Generally,
22 transfers between, for example, disk devices and random-access memories are in relatively large
23 blocks, and transfers between the random-access memories and cache memories are in somewhat
24 smaller "cache lines." In both cases, information is copied to the random-access memory and cache

2

-2-

1 memory on an "as needed" basis, that is, when the processor determines that it needs particular
2 information in its processing, it will enable blocks or cache lines which contain information to be
3 copied to the respective next faster information storage level in the memory hierarchy. Certain
4 prediction methodologies have been developed to attempt to predict the whether a processor will
5 need information for processing before it (that is, the processor) actually needs the information, and
6 to enable the information to be copied to the respective next faster information storage level.
7 However, generally at some point in the processing operations, the processor will determine that
8 information required for processing is not available in the faster information storage level, that is, a
9 "read miss" will occur, and it (that is, the processor) will need to delay its processing operations until
10 the information is available. Generally, the rate at which read misses will occur with storage
11 element(s) at a particular level in the hierarchy will be related to the storage capacity of the storage
12 element(s) at the particular level, as well as the pattern with which the processor accesses the
13 information in the respective storage level. In any case, to enhance the processing efficiency of a
14 digital data processing system, it is generally helpful to be able to assess the effect of changing the
15 capacity of the memory element(s) at a particular level in the memory hierarchy on the rate of read
16 misses at the particular level.

17 SUMMARY OF THE INVENTION

18 The invention provides a new and improved system and method for providing a prediction of
19 the operational effectiveness of a cache memory of a particular size in a digital data processing
20 system. The invention facilitates the efficient determination of the likely effectiveness of the cache
21 memory for various cache memory sizes, based on a prediction of the likely cache miss rate, the
22 prediction being based on operational statistics which are gathered during actual use of the cache
23 memory over one or more time periods, and based on a variety of cache management methodologies.
24 Based on the prediction, the operator or the system can facilitate increasing or decreasing the size of
25 the cache memory, or maintaining the cache memory at its then-current size.

-3-

1 The system determines the cache memory's read miss rate from statistics that are collected
2 during use of the cache memory over an arbitrary time interval, including statistics concerning the file
3 information retrieval activity and the extent of activity per unit time for system. Based on the
4 statistics, equations, which are based on the respective cache memory management methodology,
5 including the FIFO (first-in/first-out) methodology or the LRU (least-recently used) methodology,
6 used in managing the cache memory, are solved to generate a prediction of the cache miss rate for
7 a particular cache memory size, which may be larger or smaller than the current cache memory size,
8 and for the particular cache memory management methodology. The system can repeat this a number
9 of times over respective time intervals to determine corresponding predictions based on the cache
10 memory utilization for respective sets of statistics determined during each time interval. Thereafter,
11 the system or an operator can effect a change in the cache memory size based on the cache miss rate
12 predictions.

13 BRIEF DESCRIPTION OF THE DRAWINGS

14 This invention is pointed out with particularity in the appended claims. The above and further
15 advantages of this invention may be better understood by referring to the following description taken
16 in conjunction with the accompanying drawings, in which:

17 FIG. 1 is a functional diagram of a system for assessing the operational effectiveness of a
18 cache memory in a digital data processing system;

19 FIG. 2 is a functional block diagram of an illustrative digital data processing system with
20 which the cache assessment system depicted in FIG. 1 can be used; and

21 FIG. 3 is a flow diagram illustrating operations performed by the cache assessment system
22 depicted in FIG. 1.

23 DETAILED DESCRIPTION OF AN ILLUSTRATIVE EMBODIMENT

if

1 The invention provides a new system 10 for generating an assessment as to the operational
2 effectiveness of a cache memory operating in a digital data processing system. In an illustrative
3 embodiment, the system 10 includes a suitably programmed digital computer system, which will be
4 described in connection with FIG. 1. The computer system performs a number of processing
5 operations, as will be described below in connection with the flow chart depicted in FIG. 3, in
6 connection with operational statistics which are generated during operations in connection with a
7 respective cache memory to generate the operational effectiveness assessment. An illustrative digital
8 data processing system including a cache memory for which the cache assessment generating system
9 10 generates an operational assessment will be described in connection with FIG. 2.

10 With initial reference to FIG. 1, the cache assessment system 10 in one embodiment includes
11 digital computer system including a processor module 11 and operator interface elements comprising
12 operator input components such as a keyboard 12A and/or a mouse 12B (generally identified as
13 operator input element(s) 12) and an operator output element such as a video display device 13. The
14 illustrative computer system is of the conventional stored-program computer architecture. The
15 processor module 11 includes, for example, processor, memory and mass storage devices such as disk
16 and/or tape storage elements (not separately shown) which perform processing and storage
17 operations in connection with digital data provided thereto. In addition, the processor module 11 can
18 include one or more network ports which are connected to communication links which connect the
19 computer system in a computer network. The network ports enable the computer system to transmit
20 information to, and receive information from, other computer systems and other devices in the
21 network. The operator input element(s) 12 are provided to permit an operator to input information
22 for processing. The video display device 13 is provided to display output information generated by
23 the processor module 11 on a screen to the operator, including data that the operator may input for
24 processing, information that the operator may input to control processing, as well as information
25 generated during processing. The processor module 11 generate information for display by the video
26 display device 13, in one embodiment using a so-called "graphical user interface" ("GUI"). Although
27 the computer system is shown as comprising particular components, such as the keyboard 12A and

5

-5-

1 mouse 12B for receiving input information from an operator, and a video display device 13 for
2 displaying output information to the operator, it will be appreciated that the computer system may
3 include a variety of components in addition to or instead of those depicted in FIG. 1.

4 As noted above, the cache assessment system 10 constructed in accordance with the invention
5 generates an assessment as to the operational effectiveness of a cache memory in a digital data
6 processing system. An illustrative such digital data processing system 14 is depicted in FIG. 2. With
7 reference to FIG. 2, digital data processing system 14 includes a plurality of host computers 15(1)
8 through 15(N) (generally identified by reference numeral 15(n)) and a digital data storage subsystem
9 16 interconnected by a common bus 17. Each host computer 15(n) may comprise, for example, a
10 personal computer, workstation, or the like which may be used by a single operator, or a multi-user
11 computer system which may be used by a number of operators. Each host computer 15(n) is
12 connected to an associated host adapter 24(n), which, in turn, is connected to bus 17. Each host
13 computer 15(n) may control its associated host adapter 24(n) to perform a retrieval operation, in
14 which the host adapter 24(n) initiates retrieval of computer programs and digital data (generally,
15 "information") from the digital data storage subsystem 16 for use by the host computer 15(n) in its
16 processing operations. In addition, the host computer 15(n) may control its associated host adapter
17 24(n) to perform a storage operation in which the host adapter 24(n) initiates storage of processed
18 data in the digital data storage subsystem 16. Generally, retrieval operations and storage operations
19 in connection with the digital data storage subsystem 16 will collectively be referred to as "access
20 operations."

21 In connection with both retrieval and storage operations, the host adapter 15(n) will transfer
22 access operation command information, together with processed data to be stored during a storage
23 operation, over the bus 17. Access to the bus 17 is controlled by bus access control circuitry which,
24 in one embodiment, is integrated in the respective host adapters 24(n). The bus access control
25 circuitry arbitrates among devices connected to the bus 17 which require access to the bus 17. In
26 controlling access to the bus 17, the bus access control circuitry may use any of a number of known
27 bus access arbitration techniques.

6

1 The digital data storage subsystem 16 in one embodiment is generally similar to the digital
2 data storage subsystem described in U. S. Patent No. 5,206,939, entitled System And Method For
3 Disk Mapping And Data Retrieval, issued April 27, 1993 to Moshe Yanai, et al (hereinafter, "the '939
4 patent"). As shown in FIG. 2, the digital data storage subsystem 16 includes a plurality of digital data
5 stores 20(1) through 20(M) (generally identified by reference numeral 20(m)), each of which is also
6 connected to bus 17. Each of the data stores 20(m) stores information, including programs and data,
7 which may be accessed by the host computers 15(n) for processing, as well as processed data
8 provided to the digital data storage subsystem 16 by the host computers 15(n).

9 Each data store 20(m), in turn, includes a storage controller 21(m) and one or more storage
10 devices generally identified by reference numeral 22. The storage devices 22 may comprise any of
11 the conventional magnetic disk and tape storage devices, as well as optical disk storage devices and
12 CD-ROM devices from which information may be retrieved. Each storage controller 21(m) connects
13 to bus 17 and controls the storage of information which it receives thereover in the storage devices
14 connected thereto. In addition, each storage controller 21(m) controls the retrieval of information
15 from the storage devices 22 which are connected thereto for transmission over bus 17, and in one
16 embodiment includes bus access control circuitry for controlling access to bus 17.

17 The digital data storage subsystem 16 also includes a common memory subsystem 30 for
18 caching information during an access operation and event status information providing selected status
19 information concerning the status of the host computers 15(n) and the data stores 20(m) at certain
20 points in their operations. The caching of event status information by the common memory
21 subsystem 30 is described in detail in U. S. Patent Appn. Ser. No. 08/532,240 filed September 25,
22 1995, in the name of Eli Shagam, et al., and entitled Digital Computer System Including Common
23 Event Log For Logging Event Information Generated By A Plurality of Devices (Atty. Docket No.
24 95-034) assigned to the assignee of the present invention and incorporated herein by reference. The
25 information cached by the common memory subsystem 30 during an access operation includes data
26 provided by a host computer 15(n) to be stored on a data store 20(m) during a storage operation, as

7

-7-

1 well as data provided by a data store 20(m) to be retrieved by a host computer 15(n) during a
2 retrieval operation.

3 The common memory subsystem 30 effectively operates as a cache to cache information
4 transferred between the host computers 15(n) and the data stores 20(m) during an access operation.

5 The common memory subsystem 30 includes a cache memory 31, a cache index directory 32 and a
6 cache manager 33, which are generally described in U. S. Pat. Appn. Ser. No. 07/893,509 filed June
7 4, 1995, in the name of Moshe Yanai, et al., entitled "System And Method For Dynamically
8 Controlling Cache Management," and U. S. Pat. Appn. Ser. No. _____, filed September 2,

9 ~~1995, in the name of Eli Shagam, and entitled Average Flow Through Time In Cache (Atty. Docket~~
10 ~~No. 95-032) (hereinafter referred to as the "Shagam application"), both of which are assigned to the~~

11 assignee of the present invention and incorporated herein by reference. The cache memory 31
12 operates as a buffer in connection with storage and retrieval operations, in particular buffering
13 information received from the data stores 20(m) requested by the host computers 15(n) for
14 processing, as well as information received from the host computers 15(n) to be transferred to the
15 storage devices for storage.

16 In operation, when a host computer 15(n) wishes to retrieve information from the storage
17 subsystem 16, it initially enables its host adapter 24(n), in particular a cache manager 25(n) associated
18 with the host adapter 24(n), to determine whether the information to be retrieved is in the cache
19 memory 31. If the information to be retrieved is in the cache memory 31, that is, if a "read hit" occurs
20 in connection with the cache memory 31, the cache manager 25(n) will retrieve the information from
21 the cache memory 31 and transfer it to the host computer 15(n). On the other hand, if the
22 information to be retrieved is not in the cache memory 31, that is, if a "read miss" occurs in
23 connection with the cache memory 31, the cache manager 25(n) will enable a cache manager 23(m)
24 associated with the data store 20(m) whose storage device 22 which contains the information to be
25 retrieved to perform a "staging operation" to transfer a portion of a file containing the information
26 to be retrieved from the storage device 22 to the cache memory 31. After the portion of the file has
27 been transferred from the storage device 22 to the cache memory 31 during the staging operation,

8

-8-

1 the data store's cache manager 23(m) will notify the host adapter's cache manager 25(n), after which
2 the host adapter's cache manager 25(n) can retrieve the information from the cache memory 31.

3 Similar operations are performed in connection with a storage operation, in which information
4 in a file is updated. In particular, during a storage operation, the host adapter 24(n) will enable its
5 cache manager 25(n) to initially determine whether at least the portion of the file to be updated is in
6 the cache memory 31. If the portion of the file to be updated is in the cache memory, that is, if a
7 "write hit" occurs in connection with the cache memory 31, the cache manager 25(n) will store the
8 updated information in the cache memory 31. On the other hand, if the cache manager 25(n)
9 determines that the portion of the file to be updated is not in the cache memory 31, that is, if a "write
10 miss" occurs in connection with the cache memory 31, the cache manager 25(n) will enable the cache
11 manager 23(m) associated with the data store 20(m) whose storage device 22 which contains the
12 portion of the file to be updated to perform a "staging operation" to transfer the data from the storage
13 device 22 to the cache memory 31. After the portion of the file has been transferred from the storage
14 device 22 to the cache memory 31 during the staging operation, the data store's cache manager 23(m)
15 will notify the host adapter's cache manager 25(n), after which the host adapter's cache manager 25(n)
16 can store the updated information in the cache memory 31.

17 It will be appreciated that the efficiency of the digital data processing system 14 will generally
18 be enhanced if the rate at which, in particular, read misses in connection with cache memory 31 can
19 be reduced. The rate at which read misses occur is of importance, relative to write misses, since read
20 misses can slow down the rate at which the host computers 15(n) will be able to obtain information
21 for processing, whereas write misses will just slow down the rate at which updated information is
22 stored in the storage subsystem 16. The cache assessment system 10 provides an assessment as to
23 the effectiveness of the common memory subsystem 30 in operating as a cache, and in particular
24 generates an assessment as to the changes in read misses which may occur if the cache memory 31
25 is increased or decreased in size.

26 The cache assessment system 10 determines the cache memory's read miss rate from statistics
27 that are collected during operation of the digital data processing system 14, including statistics

9

-9-

concerning the file information retrieval activity and the extent of activity per unit time for digital data processing system 14. The file access activity, which will be represented by A_i , measures the number of times the host computers 15(n) issued requests to retrieve information from the particular "i-th" file, including those retrieval requests for which the information was already in the cache memory 31 and those retrieval requests for which staging operations were required to transfer the information from the storage device 22 to the cache memory 31. The extent of activity, which will be represented by E_i , refers to the amount of the respective "i-th" file which is active. Both statistics A_i and E_i are gathered over an arbitrary time interval, and represent the activity and extent of activity over the particular time interval. Assuming that

(a) the activity is spread uniformly over a particular extent, which can occur if the host computers 15(n) randomly issue retrieval requests for information from the respective extent;

(b) the digital data processing system 14 caches all read misses in cache memory 31, that is, during staging operations in connection with information from a file following determinations by the respective host adapters 24(n) that the information was not already in the cache memory 31; and

(c) the cache memory 31 is managed on a first-in first-out (FIFO) basis, that is, information is removed from the cache in the order in which it is loaded into the cache (generally, if the cache memory 31 is large, other cache management methodologies, such as the "least-recently used" methodology, will approximate FIFO management)

then, if P_i represents the percentage of the cache memory 31 which is occupied by information from a file "i," at each point in time the sum of the percentages of all of the files "i" that are stored in the storage devices 22 of storage subsystem 16 which are cached in the cache memory 31 equals one, that is,

$$\sum_i P_i = 1 \quad (1).$$

23

T100x

10

-10-

Based on assumptions (b) and (c) above, that is, based on the assumption that the digital data processing system 14 caches all read misses in the cache memory 31, stores the staged information retrieved from a file stored on a storage device caches all of the read misses in the cache memory 31, and based further on the assumption that the cache memory 31 is managed on a FIFO basis, then the percentage of the "i-th" file that is cached in the cache memory 31 at any point in time corresponds to the ratio between the number of read misses in connection with the "i-th" file and the total number of read misses, that is,

$$P_i = \frac{M_i}{M} \quad (2),$$

where M_i represents the number of read misses in connection with file "i" per unit time and "M" represents the total number of read misses which occur in the digital data processing system 14 per unit time. At any point in time, the portion of the "i-th" file which is cached in the cache memory 31 corresponds to $\frac{P_i S}{E_i}$ and so the portion of the file which is not cached in the cache memory 31

corresponds to $1 - \frac{P_i S}{E_i}$. Based on assumption (a) above, that is, that activity rate A_i in connection

with a file is spread uniformly over the file, the number of read misses in connection with file "i" per unit time, M_i , in turn, is related to the portion of the file which is not in the cache memory 31 at any point in time, times the activity rate A_i , or

$$M_i = \left(1 - \frac{P_i S}{E_i} \right) A_i \quad (3).$$

-11-

1 Combining equations (2) and (3), the multiplicative product of the P_i , the percentage of the
 2 cache memory 31 which is occupied by information from the "i-th" file over each time interval, times
 3 the total number of read misses over the time interval corresponds to

4
$$P_i M = M_i = \left(1 - \frac{P_i S}{E_i} \right) A_i \quad (4).$$

T(12)X

5 Rearranging equation (4),

6
$$P_i M + P_i \frac{S A_i}{E_i} = A_i \quad (5)$$

7 and solving equation (5) for P_i ,

8
$$P_i = \frac{A_i}{M + \frac{S A_i}{E_i}} \quad (6).$$

T(12)X

9 Combining equation (6) and equation (1),

-12-

$$1 = \sum_i P_i = \sum_i \frac{A_i}{M + \frac{SA_i}{E_i}} \quad (7)$$

which is a function of "M," the total number of read misses per time interval, the size "S" of the cache and the statistics A_i and E_i which are generated as described above.

Equation (7) effectively gives a prediction as to the number of read misses in connection with cache memory 31 as a function of the size "S" of the cache memory 31, for the particular types of processing operations which are performed by the digital data processing system 14 during the time interval over which the statistics A_i and E_i were collected. Accordingly, for any particular size "S" as selected by the operator of the cache assessment system 10, equation (7) can be solved for "M" to provide a prediction as to the effect on the number of read misses which would occur for the particular cache memory size S. That is, for any particular value of "S," the size of the cache memory 31, the number of read misses can be predicted by providing that value in equation (7) and solving for "M."

Equation (7) can be solved for "M" in a number of ways. Equation (7) can be efficiently solved for "M" using a binary search arrangement, as will be evident from the following. First, it will be recognized that, generally, the sum in the right-hand side of equation (7) is a function of "M" and "S," that is,

$$f(M, S) = \sum_i \frac{A_i}{M + \frac{SA_i}{E_i}} \quad (8).$$

-13-

is monotonically decreasing for positive values of "M" and "S." Differentiating equation (8) with respect to "M,"

$$f'(M, S) = \sum_i \frac{-A_i}{\left(M + \frac{SA_i}{E_i}\right)^2} \quad (9)$$

where $f(M, S)$ represents the derivative of $f(M, S)$ with respect to M . Since A_i , E_i , M and S are all positive numbers, it will be appreciated that the derivative is negative for all positive values of "M" and "S", and so $f(M, S)$ as shown in equation (8) is monotonically decreasing for all values of "S." Thus, for each value of "S" there will be one real value for "M." In addition, it will be recognized that the value of "M" will fall in the interval

$$0 \leq M \leq \sum_i A_i = A \quad (10)$$

(where "A" represents the total activity over the time interval), since the number of read misses in a time interval cannot be negative and cannot exceed the total activity over the time interval. Any conventional methodology can be used to determine or approximate the solution to equation (7), including, for example, a conventional binary search methodology over the interval defined by equation (10).

Using equations (7) and (10), an operator of the cache assessment system 10 can assess the effectiveness, if any, of increasing or decreasing the size of the cache memory 31 from its current size, based on the particular types of processing activities performed during the time interval over which

-14-

1 the statistics A_i and E_i were collected. Using equations (7) and (10) over a plurality of time intervals,
2 during which the mixture of types of processing operations will likely change, the operator can
3 determine the effectiveness, if any, of increasing or decreasing the size of the cache memory 31 for
4 the diverse mixtures of processing operations which the digital data processing apparatus 14 may be
5 called upon to perform.

6 As noted above, the model described in connection with equations (1) through (9)
7 encompasses assumptions (a) and (b) above, namely, that the activity in connection with an accessed
8 file is spread uniformly over the respective file (assumption (a)), and that the digital data processing
9 system 14 caches all read misses in cache memory 31 (assumption (b)). For at least information to
10 be processed (in contrast to program instructions which may be stored in the storage subsystem 16
11 and retrieved by the host computers 15(n)), generally assumption (b) is correct, that is, generally such
12 information retrieved from the storage devices in connection with a staging output following a read
13 miss will be cached in the cache memory 31. However, activity in connection with an accessed file
14 does not need to be uniformly distributed uniformly over the entire file, and so assumption (a) may
15 not be correct.

16 To extend the methodology to a model in which assumption (a) is eliminated, it will be
17 assumed that each file may comprise one or more relatively short, non-overlapping, and possibly
18 variable-length "packets." When the an access request is issued to particular packet is referenced,
19 the packet may be re-referenced within a relatively short period of time. As an illustration which will
20 assist in clarifying these assumptions, in a file containing banking information for customers of a bank,
21 each of the packets may comprise information concerning a particular account or customer. In that
22 case, the packets may have differing lengths, based on the amount of banking the customer has done
23 with the particular bank. In addition, typically the packets will be accessed when, for example, the
24 customer's account is updated by the bank or when the customer calls for information concerning the
25 account, and will otherwise not be accessed. In that case, when a packet is accessed, a number of
26 references may be made to the packet within a short time, that is, while the bank is performing the
27 update operation or during the call. In this extended methodology, the activity and extent statistics

15

-15-

A_i and E_i , are still determined on a file basis; however, a further statistic is generated, namely, a reference average R_i , which indicates the average number of times a packet is referenced during each time interval, that is, the number of times a host computer 15(n) retrieves information from the packet. (A methodology for determining the value of R_i will be described below.) It will be appreciated that the reference average R_i includes the initial reference, which would give rise to a cache miss. Under these assumptions, equation (4) becomes

$$P_i M = M_i = \left[\left(1 - \frac{P_i S}{E_i} \right) + (R_i - 1 - H(i, S)) \right] \frac{A_i}{R_i} \quad (11)$$

where $H(i, S)$ corresponds to the number of cache hits per packet. (A methodology for determining the value of $H(i, S)$ will be described below.) In equation 11,

(a) the term " $\left(1 - \frac{P_i S}{E_i} \right)$ ", as in equation (4), is an indication of the amount of the amount

of the "i-th" file that is not in the cache, and so $\left(1 - \frac{P_i S}{E_i} \right) \frac{A_i}{R_i}$ is an indication of the number of

cache misses on the first access request in connection with a packet; and

(b) the term " $(R_i - 1 - H(i, S)) \frac{A_i}{R_i}$ " is an indication of the number of cache misses on

subsequent references in connection with a packet.

Rearranging equation (11) in the same manner as equation (4) above,

-16-

$$P_i \left(M + \frac{S \left(\frac{A_i}{R_i} \right)}{E_i} \right) = \frac{A_i}{R_i} (R_i - 1 - H(i, S)) \quad (12)$$

2 and solving equation (12) for P_i ,

$$P_i = \frac{\frac{A_i}{R_i} (R_i - 1 - H(i, S))}{M + \frac{S \left(\frac{A_i}{R_i} \right)}{E_i}} \quad (13)$$

4 Since, as above in connection with equation (1), it is assumed that the cache memory 31 is fully
5 populated with information from the data stores 20(m),

$$1 = \sum_i P_i = \sum_i \frac{\frac{A_i}{R_i} (R_i - 1 - H(i, S))}{M + \frac{S \left(\frac{A_i}{R_i} \right)}{E_i}} \quad (14),$$

-17-

1 which, as in equation (7), provides the cache miss ratio "M" as a function of the cache size "S" and
 2 the statistics developed over the time interval.

3 Using a methodology similar to that described above in connection with equation (7), equation
 4 14 can be solved in the following manner. From the following, it will be clear that ^{equation} ~~equation~~ (14)
 5 provides one real solution of "M" for each value of "S." First, it will be recognized that, generally,
 6 the sum in the right-hand side of equation (14) is a function of "M" and "S," that is,

7

$$f(M, S) = \sum_i \frac{\frac{A_i}{R_i} (R_i - 1 - H(i, S))}{M + \frac{S \left(\frac{A_i}{R_i} \right)}{E_i}} \quad (15)$$

8 is monotonically decreasing for positive values of "M" and "S." Differentiating equation (15) with
 9 respect to "M,"

10

$$f'(M, S) = \sum_i \frac{-\frac{A_i}{R_i} (R_i - 1 - H(i, S))}{\left(M + S \frac{\left(\frac{A_i}{R_i} \right)}{E_i} \right)^2} \quad (16)$$

-18-

where $f(M,S)$ represents the derivative of $f(M,S)$ with respect to M . Since A_i , E_i , M and S are all positive numbers, it will be appreciated that the derivative is negative for all positive values of " M " and " S ", and so $f(M,S)$ as shown in equation (8) is monotonically decreasing for all values of " S ." Thus, for each value of " S " there will be one real value for " M ." In addition, it will be recognized that the value of " M " will fall in the interval

$$0 \leq M \leq \sum_i A_i = A \quad (17)$$

(where " A " represents the total activity over the time interval), since the number of read misses in a time interval cannot be negative and cannot exceed the total activity over the time interval. Any conventional methodology can be used to determine or approximate the solution to equation (14), including, for example, a conventional binary search methodology over the interval defined by equation (17).

Using equations (14) and (17), an operator of the cache assessment system 10 can assess the effectiveness, if any, of increasing or decreasing the size of the cache memory 31 from its current size, based on the particular types of processing activities performed during the time interval over which the statistics A_i and E_i were collected. Using equations (14) and (17) over a plurality of time intervals, during which the mixture of types of processing operations will likely change, the operator can determine the effectiveness, if any, of increasing or decreasing the size of the cache memory 31 for the diverse mixtures of processing operations which the digital data processing apparatus 14 may be called upon to perform.

As noted above, the methodologies described above in connection with equations 1 through 10 and 11 through 17 both assume that the cache memory 31 is operated using a FIFO cache operating methodology (assumption (c) above). While the FIFO methodology can be useful in approximating conditions in cache memories which use other methodologies, particularly if the cache memories are relatively large, a more accurate methodology for use with a cache memory which is

-19-

operated using the "LRU" (least recently used) cache management methodology, will be described in connection with equations (18) through (21) below. As is usual in the LRU cache management methodology, when a cache miss occurs and information is staged into the cache memory 31, the information that is replaced during the staging operation is the information that was least recently accessed by, for example, a host adapter 24(n). To maintain the LRU ordering, when information in the cache memory 31 is accessed, the accessed information's position in the LRU ordering is promoted to the top of the LRU ordering. Thus, since a time interval "I," there are M/I cache slots in the cache memory 31 replaced ^{during} the time interval, the amount of time that an extent will remain in the cache memory 31 after the last time the extent was referenced will correspond to the total size of the cache memory 31, "S," divided by the number of cache slots that are replaced during a time interval "M/I." That is, the amount of time that an extent will remain in the cache memory 31 after it was last reference corresponds to S/(M/I). Accordingly, if the time interval between the first time the extent is accessed and the last time the extent is accessed is "T_i," then the total time that the extent will remain in the cache memory 31 is $T_i + \frac{S}{\left(\frac{M}{I}\right)} = T_i + \frac{SI}{M}$. Thus, the percentage of

the "i-th" file in the cache memory 31 at any point in time (reference equation (2)) above corresponds to

$$P_i = \frac{M_i \left(T_i + \frac{SI}{M} \right)}{SI} \quad (18).$$

If the cache memory 31 is large, then

-20-

$$M_i = \left(1 - \frac{P_i S}{E_i} \right) \frac{A_i}{R_i} \quad (19)$$

analogous to equations (3) and (11). From equations (1), (18) and (19)

$$1 = \sum_i P_i = \sum_i \frac{\left(1 - \frac{P_i S}{E_i} \right) \frac{A_i}{R_i} \left(T_i + \frac{SI}{M} \right)}{SI} \quad (20).$$

Rearranging equaiton (20),

$$1 = \sum_i P_i = \sum_i \frac{\left(\frac{A_i T_i}{R_i} \right) M + \left(\frac{A_i SI}{R_i} \right)}{\left(SI + \frac{SA_i T_i}{E_i R_i} \right) M + \left(\frac{A_i I}{E_i R_i} \right) S^{\parallel}} \quad (21),$$

which has a form similar to equations (7) and (14), and can be solved in a similar manner using a binary search technique. Using equations (14) and (17), an operator of the cache assessment system can assess the effectiveness, if any, of increasing or decreasing the size of the cache memory

-21-

1 from its current size, based on the particular types of processing activities performed during the time
2 interval over which the statistics A_i and E_i were collected. Thus, using equation (21), along with
3 equations (10) and (17) (which define the effective range for "M," the number of cache misses) over
4 a plurality of time intervals, during which the mixture of types of processing operations will likely
5 change, the operator can determine the effectiveness, if any, of increasing or decreasing the size of
6 the cache memory 31 for the diverse mixtures of processing operations which the digital data
7 processing apparatus 14 may be called upon to perform.

8 As noted above, the second and third methodologies, described above in connection with
9 equations 11 through 17 and 18 through 21, respectively, require values for variable R_i , $H(i,S)$ and
10 T_i to be determined over respective time intervals. Such values can be determined as follows. The
11 value of T_i , the times between the first and last access to a cached extent in the cache memory 31, can
12 be determined by providing a time stamp for each cache slot, which includes an initial value when an
13 extent is assigned to the cache slot, and an access value that is updated each time the cache slot is
14 accessed. In that case, when the cache slot is re-used for another extent, the value of T_i for the
15 previous extent will correspond to the difference between the access value and the initial value.

16 The reference average R_i , which indicates the average number of times a packet in a cache
17 slot is accessed during each time interval, and $H(i,S)$, the number of cache hits per packet, can be
18 determined by establishing respective tables having an entry associated with each packet, which are
19 incremented when respective packet is accessed. When the cache slot is re-used for another extent,
20 the number of times the extent was accessed while the packet was in the cache slot can be
21 determined. The respective tables can, for example, have one entry for each extent, or, alternatively,
22 may be in the form of a hash table having a plurality of entries which are hashed to identifiers for the
23 extents in a conventional manner.

24 With this background operations performed by the cache assessment system 10 will be
25 described in connection with the flowchart in FIG. 3. With reference to FIG. 3, the cache assessment
26 system 10 initially collects the operational statistics during operation of the digital data processing
27 system 14 over a selected time period (step 100). If the cache assessment system 10 uses the first

-22-

1 model, described above in connection with equations 1 through 10, it need only collect the activity
2 and extent statistics A_i and E_i described above. On the other hand, if the cache assessment system
3 10 uses the second model, it will additionally need to collect the re-reference and cache hit statistics
4 R_i and $H(i,S)$ over the time interval.

5 After the time interval has terminated, the cache assessment system 10 will apply the statistics
6 to the appropriate equation (8) or (14), depending on the selected model, for various sizes "S" of
7 cache memory 31 to generate respective predictions as to the number of read misses for each of the
8 respective cache sizes "S" (step 101). Based on the respective predictions, the cache assessment
9 system 10 or the operator can determine whether to adjust the size of the cache memory 31 (step
10 102) and, if so, initiate operations to enable the size adjustment to occur.

11 The invention provides a number of advantages. In particular, the invention provides a system
12 and method for assessing the effect of read misses in connection with a cache memory used in a
13 digital data processing system, as a function of the size of the cache memory, using statistics that are
14 generated during processing operations of the digital data processing system. Based on the
15 assessment, a determination can be made as to the utility of providing a cache memory of a particular
16 size in the digital data processing system.

17 It will be appreciated that a number of modifications may be made to the cache assessment
18 system 10 described above. For example, although the system 10 has been described in connection
19 with a particular digital data processing system, it will be appreciated that the system 10 will find
20 utility in connection with digital data processing systems of numerous architectures.

21 It will be appreciated that a system in accordance with the invention can be constructed in
22 whole or in part from special purpose hardware or a general purpose computer system, or any
23 combination thereof, any portion of which may be controlled by a suitable program. Any program
24 may in whole or in part comprise part of or be stored on the system in a conventional manner, or it
25 may in whole or in part be provided in to the system over a network or other mechanism for
26 transferring information in a conventional manner. In addition, it will be appreciated that the system

23

-23-

1 may be operated and/or otherwise controlled by means of information provided by an operator using
2 operator input elements (not shown) which may be connected directly to the system or which may
3 transfer the information to the system over a network or other mechanism for transferring information
4 in a conventional manner.

5 The foregoing description has been limited to a specific embodiment of this invention. It will
6 be apparent, however, that various variations and modifications may be made to the invention, with
7 the attainment of some or all of the advantages of the invention. It is the object of the appended
8 claims to cover these and such other variations and modifications as come within the true spirit and
9 scope of the invention.

10 What is claimed as new and desired to be secured by Letters Patent of the United States is:

11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193